LA-UR-21-28078

Title: Neural Density Estimation and Uncertainty Quantification for ChemCam Spectra

Author(s): Kontolati, Katiana
Panda, Nishant
Klein, Natalie Elizabeth
Moore, Juston Shane
Oyen, Diane Adele

Intended for: Report

Issued: 2021-08-12

Neural Density Estimation and Uncertainty Quantification for ChemCam Spectra

AML Fellow: Katiana Kontolati
Ph.D. student, Johns Hopkins University

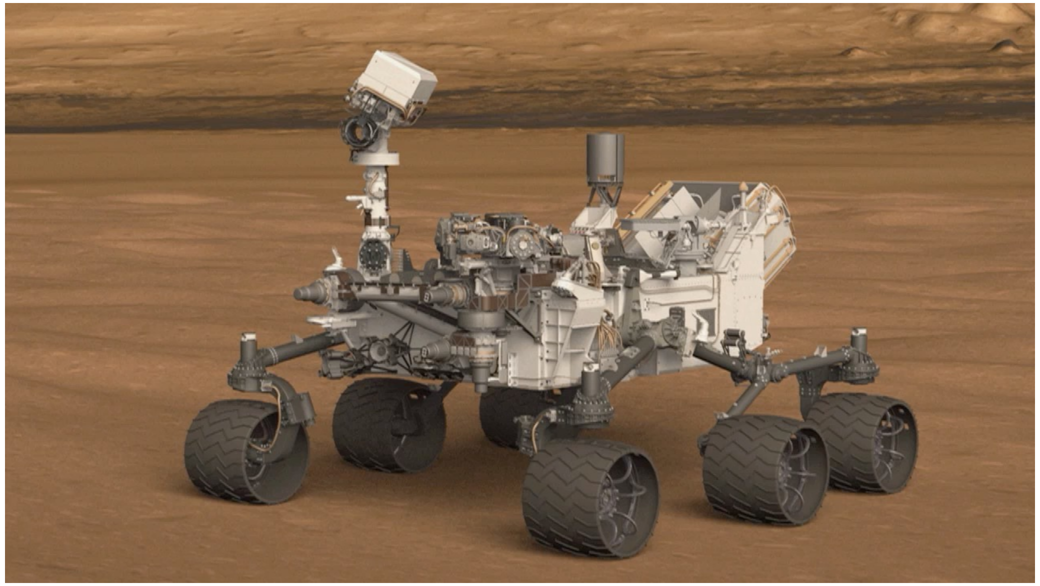
Mentors: Nishant Panda, Natalie Klein, Juston Moore, Diane Oyen

Applied Machine Learning Student Symposium
August 11, 2021

Outline

- **Background**: ChemCam data and literature review
- **Motivation**: Uncertainty quantification for ChemCam
- **Methods**: Generative modeling via normalizing flows
- Proposed approach and Results
- Discussion and Conclusions

Background: ChemCam



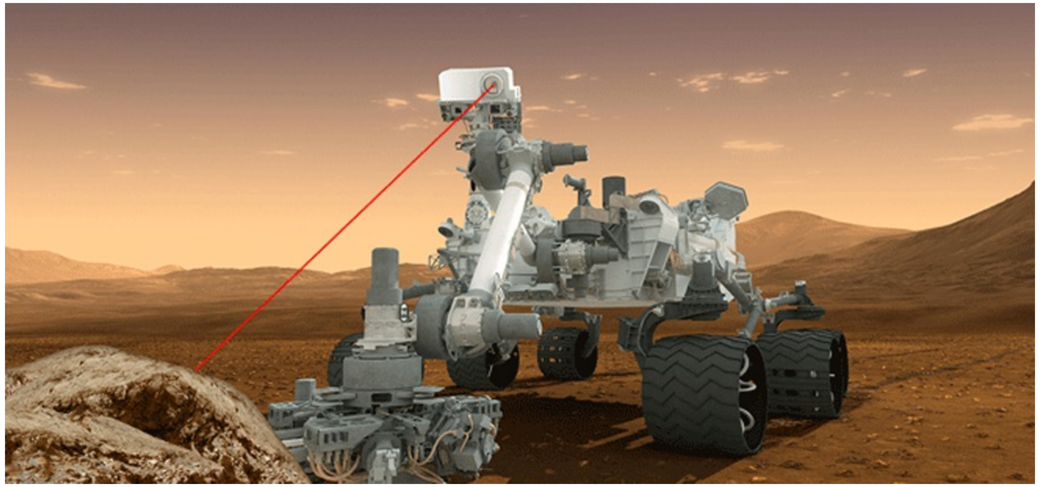
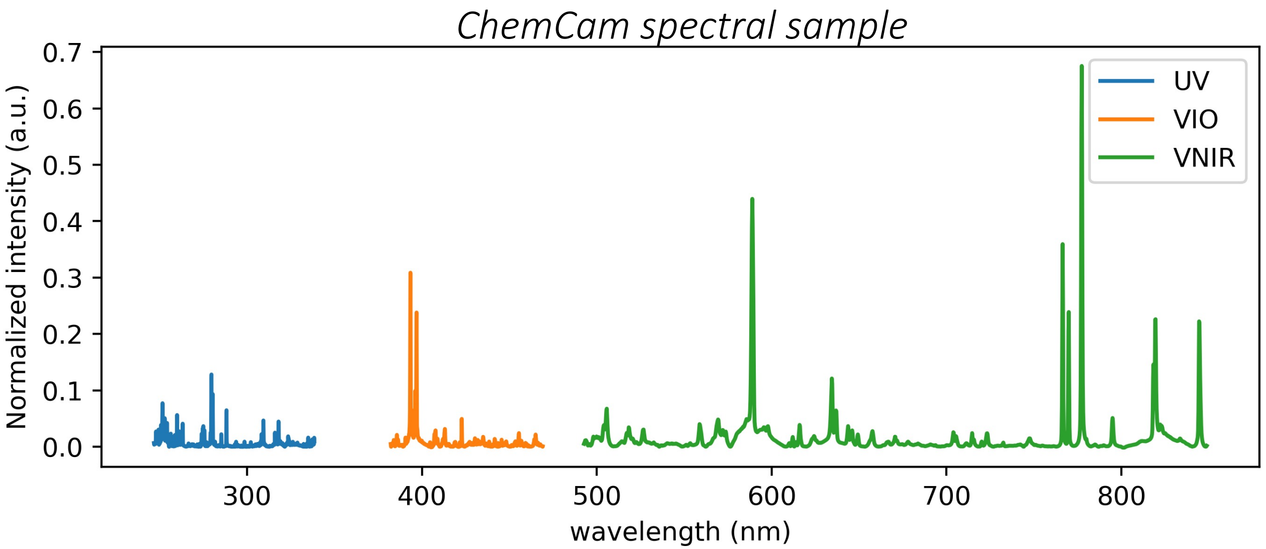


Illustration: ChemCam firing laser



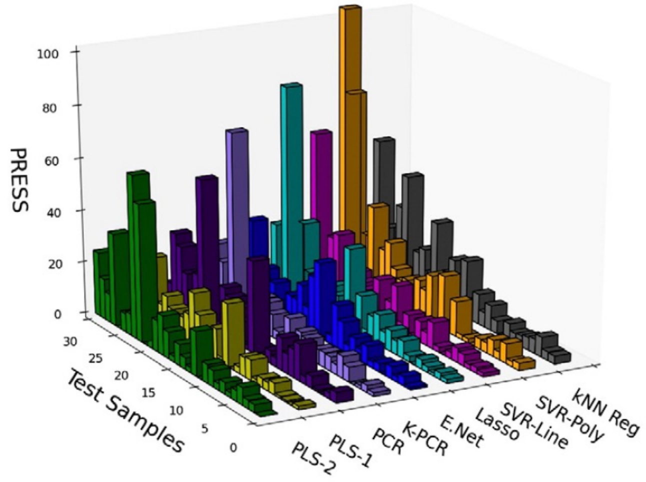
w.t.% oxide

Target	SiO2	TiO2	Al2O3	FeOT	MgO	CaO	Na2O	K2O
AGV2	59.30	1.05	16.91	6.019720	1.79	5.200	4.19	2.88

- The ChemCam instrument of Curiosity uses laser-induced breakdown spectroscopy (LIBS)
- Fires a laser at target, vaporizes rock surfaces, creating a plasma
- Three spectrographs divide the plasma light into wavelengths for chemical analysis
- The three wavelength ranges: Ultraviolet, Violet, Visible Near-Infrared

Background

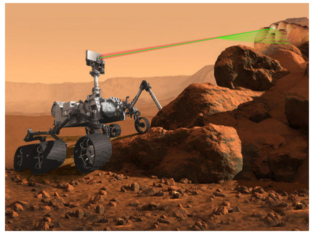
- Regression methods (SVR, PCR, CNN) have been employed for calibration (prediction of the elemental composition of samples)
- However, labeled ChemCam samples are limited

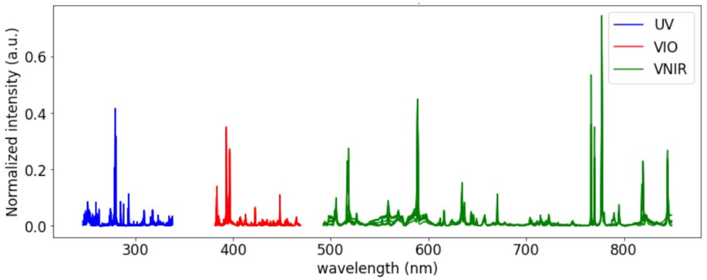


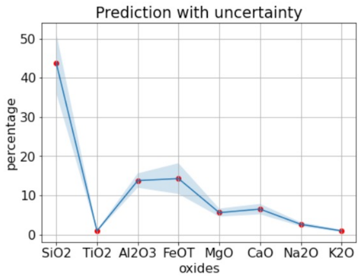
Comparison of 10 regression models
Boucher, T. F., et al., (2015). *Spectr. Acta Part B: Atomic Spectroscopy*

Motivation

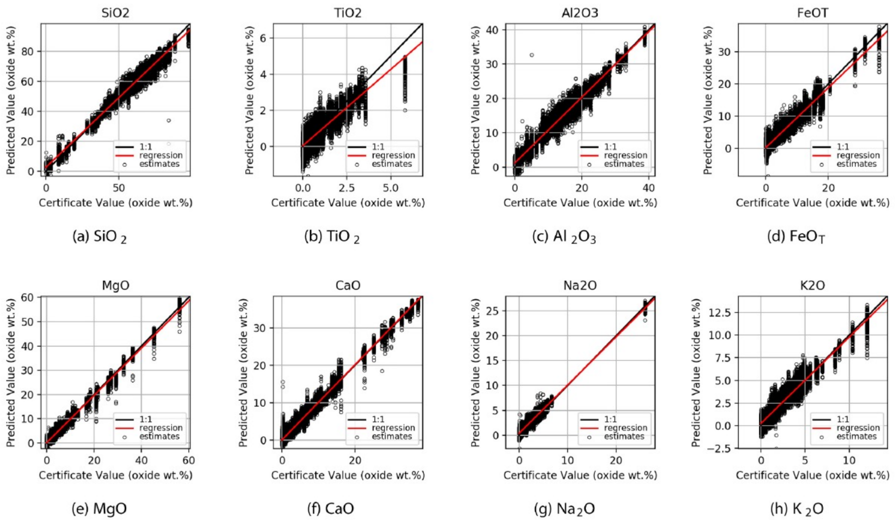
- Focus on unsupervised learning and employ *generative models* from ChemCam analysis
- Use labels (supervised) in combination to the generative model to compute uncertainties related to predictions







uncertainty propagation



Learning the chemical content of samples (regression results)
Castorena, J. et al., (2021). *Spectr. Acta Part B: Atomic Spectroscopy*

Outline

- **Background**: ChemCam data and literature review
- **Motivation**: Uncertainty quantification for ChemCam
- **Methods**: Generative modeling via normalizing flows
- Proposed approach and Results
- Discussion and Conclusions

Methods

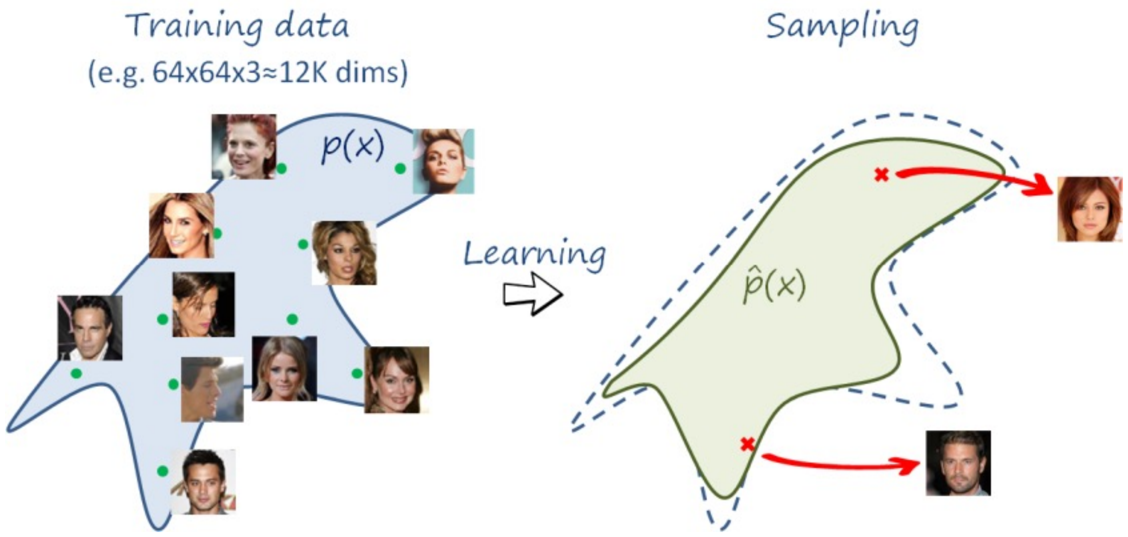
Generative modeling

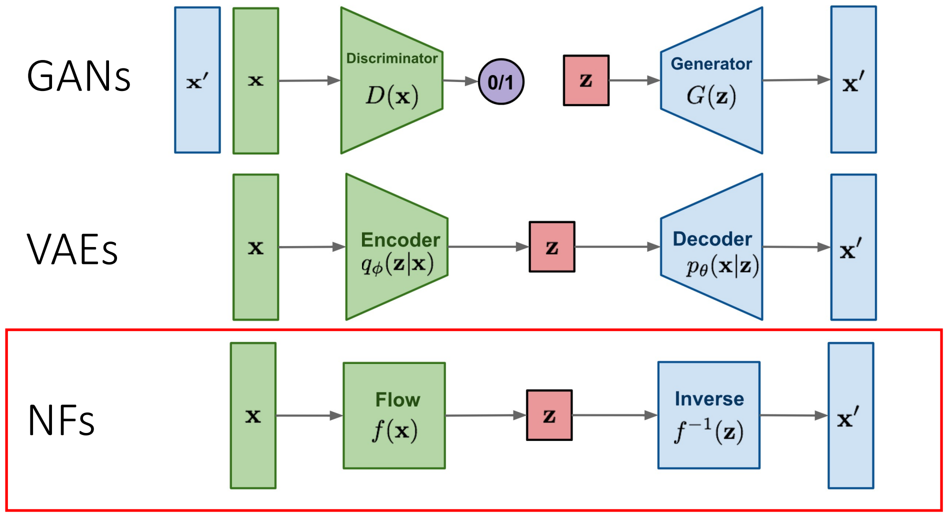
- In generative modeling, any kind of observed dataset $\mathcal{D}$, is a finite set of samples generated from an *underlying distribution*

- The goal of any generative model is to approximate this data distribution given access to the dataset $\mathcal{D}$

- If we are able to *learn* a good generative model, we can use the learned model for downstream *inference*

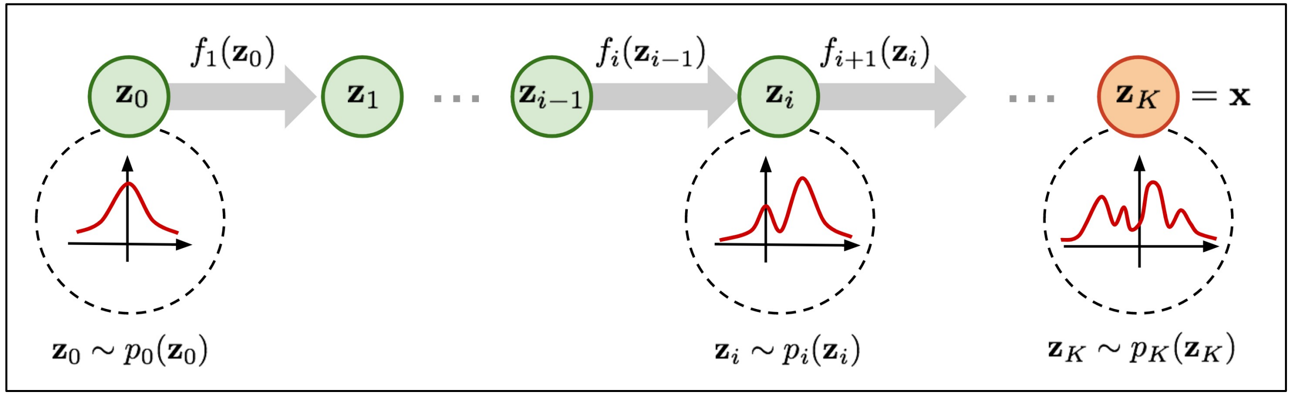
- Perform: Sampling, density estimation, detect outliers, fill in incomplete data, representation learning



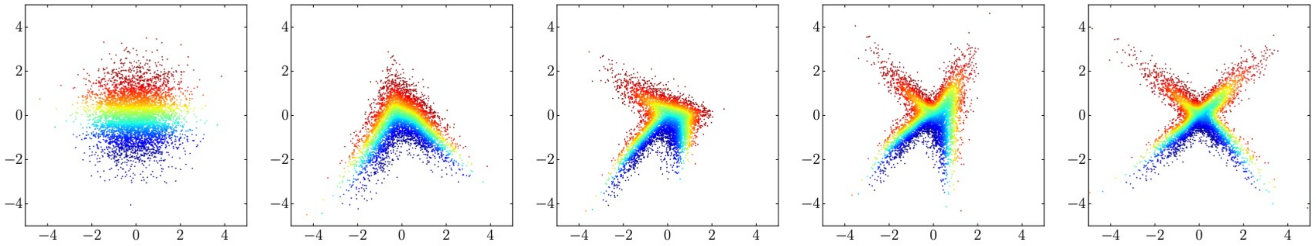


Methods: Normalizing flows (NF)

Definition: A Normalizing Flow is a transformation of a simple probability distribution (e.g., a standard normal) into a more complex distribution by a sequence of invertible and differentiable mappings.



Transformation example

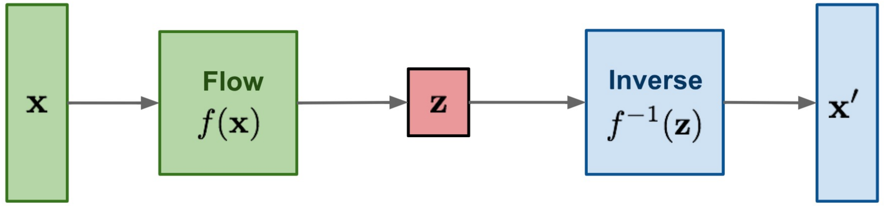


4-step flow transforming samples from a standard-normal base density to a cross-shaped target density (*Papamakarios et. al., 2021 arXiv 1912.02762*)

- Mappings $f_i(\mathbf{z}_{i-1})$ need to be computationally efficient but also expressive enough
- NF produce tractable distributions where both sampling and density evaluation can be efficient and exact
- Compared to other approaches (VAE, GAN) NF allows for exact evaluation of densities and efficient sampling
- Parallel WaveNet[1] model is currently used by Google Assistant to generate realistic speech

[1]Oord, A., et al. (2018) *ICML* (pp. 3918-3926)

Methods: Normalizing flows (NF)



- Let us consider a directed, latent-variable model over observed variables X and latent variables Z
- In a **normalizing flow model**, the mapping between Z and X, given by $f_\theta \colon \mathbb{R}^n \to \mathbb{R}^n$ is deterministic and invertible such that $X = f_\theta(Z)$ and $Z = f_\theta^{-1}(X)$
- Using change of variables, the marginal likelihood $p(\mathbf{x})$ is given by

$$p_X(\mathbf{x};\theta) = p_Z(f_\theta^{-1}(\mathbf{x}))\left|\det\left(\frac{\partial f_\theta^{-1}(\mathbf{x})}{\partial \mathbf{x}}\right)\right|$$

- "**Normalizing**": change of variables gives a normalized density after applying an invertible transformation.
- "**Flow**": invertible transformations can be composed to create more complex invertible transformations.

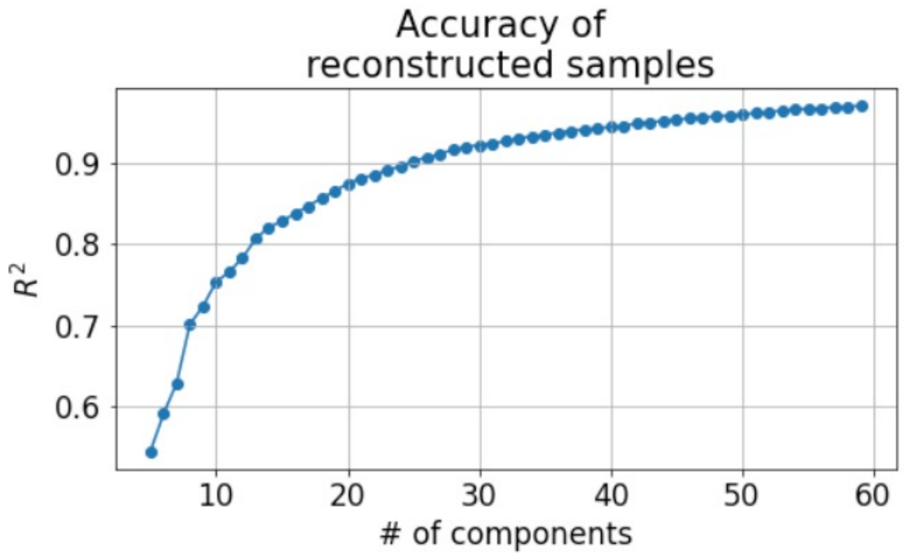
Outline

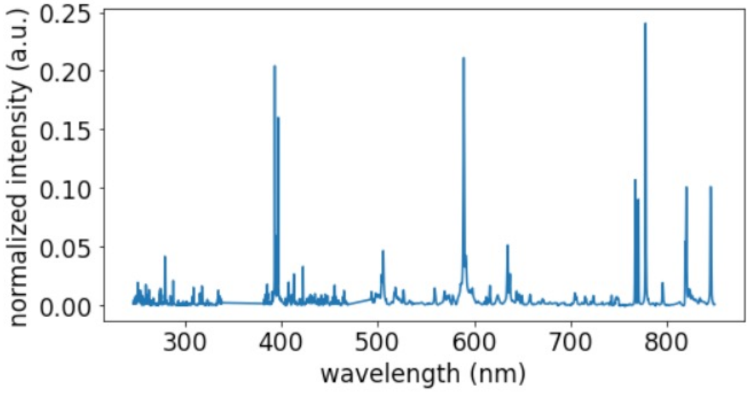
- **Background**: ChemCam data and literature review
- **Motivation**: Uncertainty quantification for ChemCam
- **Methods**: Generative modeling via normalizing flows
- Proposed approach and Results
- Discussion and Conclusions

Dimension reduction

Non-negative matrix factorization (NMF)

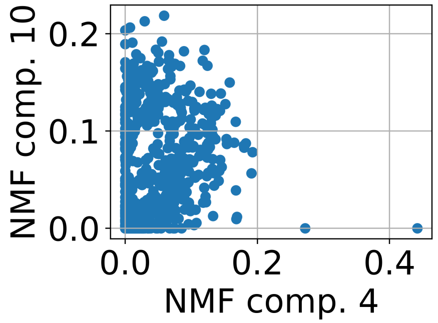
- Assume random vector $y \in \mathbb{R}^M$ where $Y = [y_1, .., y_N] \in \mathbb{R}^{NxM}$
- $M = 5606$ (original dimensionality)
- Decompose $Y \approx XV$
- $X \in \mathbb{R}_{\geq 0}^{NxL}$ and $\mathrm{V} \in \mathbb{R}_{\geq 0}^{LxM}$, $L \ll M$
- X: non-negative basis matrix, V: non-negative coefficient matrix
- Optimization: Minimize the Frobenius norm between Y and XV





original space

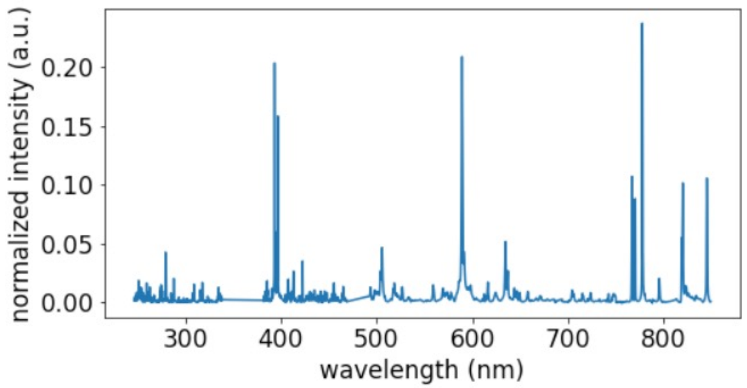
$y \in \mathbb{R}^M$



NMF

latent space

$x \in \mathbb{R}^L$



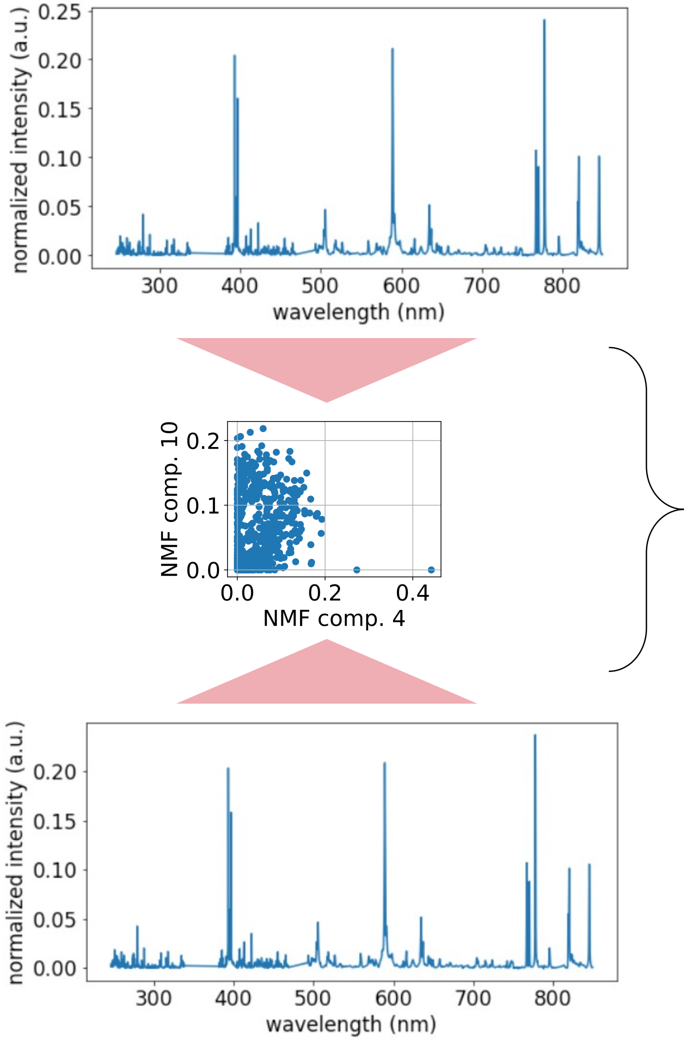
original space

$y \in \mathbb{R}^M$

Normalizing flow model

x: Spectral latent variable
z: Latent variable (~normal)

Latent space dimension: 15



Train normalizing flow model on the latent space

Real-NVP (Real-valued non-volume preserving)

Forward flow:

$$\mathbf{x}_{1:d} = \mathbf{z}_{1:d}$$

$$\mathbf{x}_{d+1:D} = \mathbf{z}_{d+1:D} \odot \exp(f_\alpha(\mathbf{z}_{1:d})) + f_\mu(\mathbf{z}_{1:d}),$$

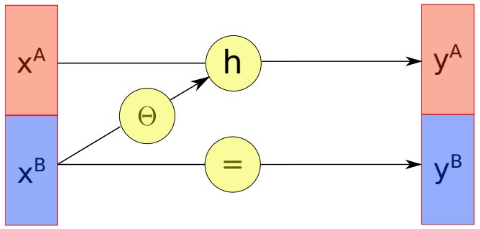
Inverse flow:

$$\mathbf{z}_{1:d} = \mathbf{x}_{1:d}$$

$$\mathbf{z}_{d+1:D} = (\mathbf{x}_{d+1:D} - f_\mu(\mathbf{x}_{1:d}) \odot \exp(-f_\alpha(\mathbf{x}_{1:d}))$$

Determinant of Jacobian:

$$\det(J) = \prod_{i=1}^{D-d} \exp(f_\alpha(\mathbf{z}_{1:d}))_i = \exp\left(\sum_{i=1}^{D-d} f_\alpha(\mathbf{z}_{1:d})_i\right)$$

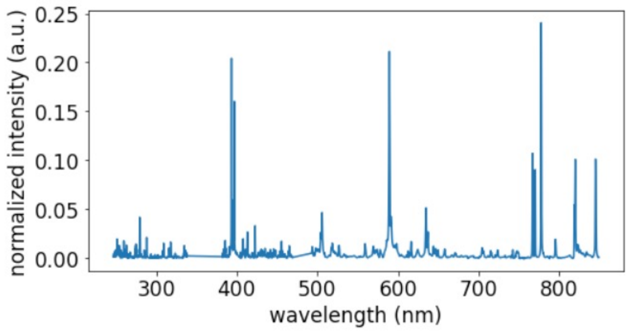


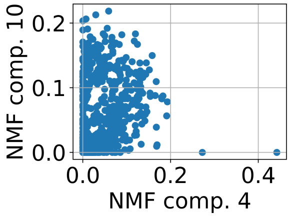
Single coupling flow architecture

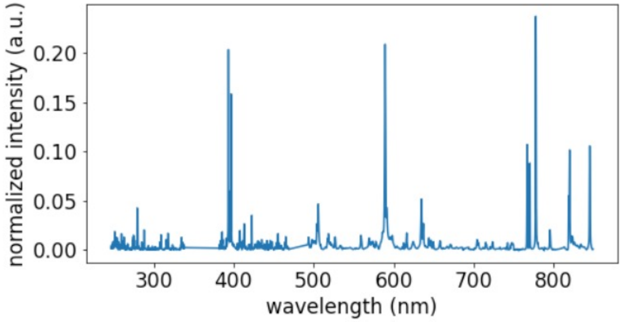
Dinh et al. (2016) Density estimation using real nvp. arXiv:1605.08803

Normalizing flow model

Latent space dimension: 15







Train normalizing flow model

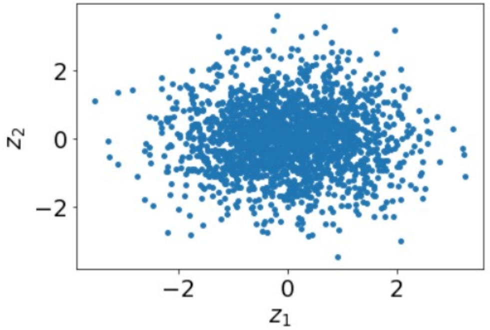
x: Spectral latent variable

z: Latent variable (~normal)

Real-NVP *Dinh et al. (2016) Density estimation using real nvp. arXiv:1605.08803*

Forward flow:
$$\mathbf{x}_{1:d} = \mathbf{z}_{1:d}$$
$$\mathbf{x}_{d+1:D} = \mathbf{z}_{d+1:D} \odot \exp(f_\alpha(\mathbf{z}_{1:d})) + f_\mu(\mathbf{z}_{1:d}),$$

Inverse flow:
$$\mathbf{z}_{1:d} = \mathbf{x}_{1:d}$$
$$\mathbf{z}_{d+1:D} = (\mathbf{x}_{d+1:D} - f_\mu(\mathbf{x}_{1:d}) \odot \exp(-f_\alpha(\mathbf{x}_{1:d}))$$

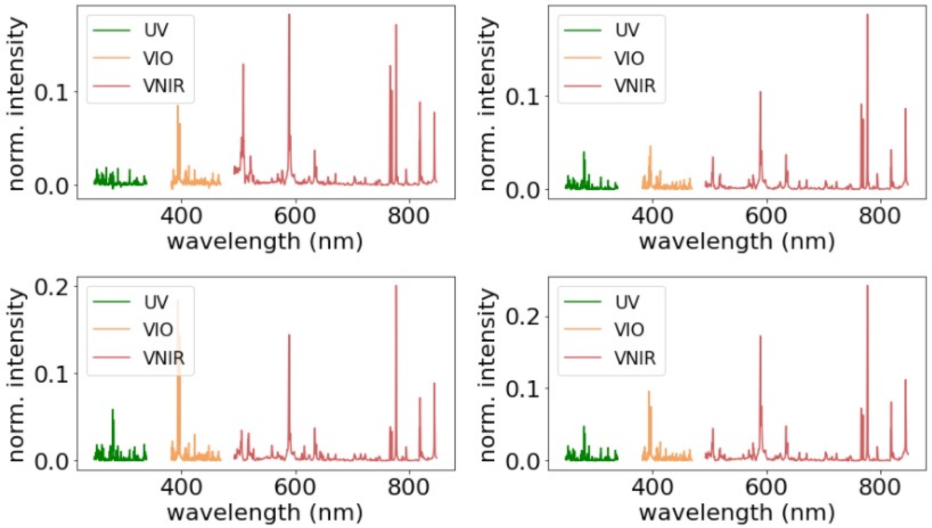


$z \sim N(0, I_{15})$

generative direction

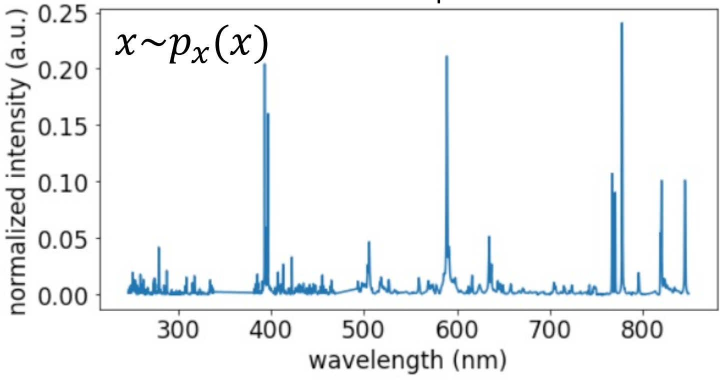
normalizing direction

$x \sim p_x(x)$



Map LIBS spectra to compositions

LIBS sample

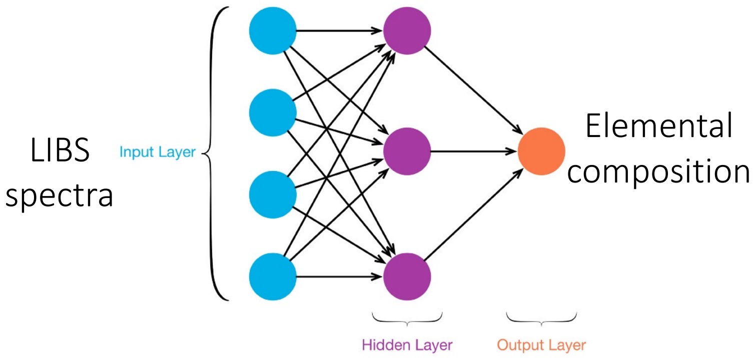


w.t.% oxides

Target	SiO2	TiO2	Al2O3	FeOT	MgO	CaO	Na2O	K2O
AGV2	59.30	1.05	16.91	6.019720	1.79	5.200	4.19	2.88
BCR-2	54.10	2.26	13.50	12.417359	3.59	7.120	3.16	1.79
BEN	38.20	2.61	10.07	11.610990	13.15	13.870	3.18	1.39
BHVO2	49.90	2.73	13.50	11.067646	7.23	11.400	2.22	0.52
BIR-1a	47.70	0.97	15.40	10.193606	9.70	13.400	1.81	0.03
...	...	...	...	...	...	...	...	...

Train a MLP model for each oxide

Architecture:



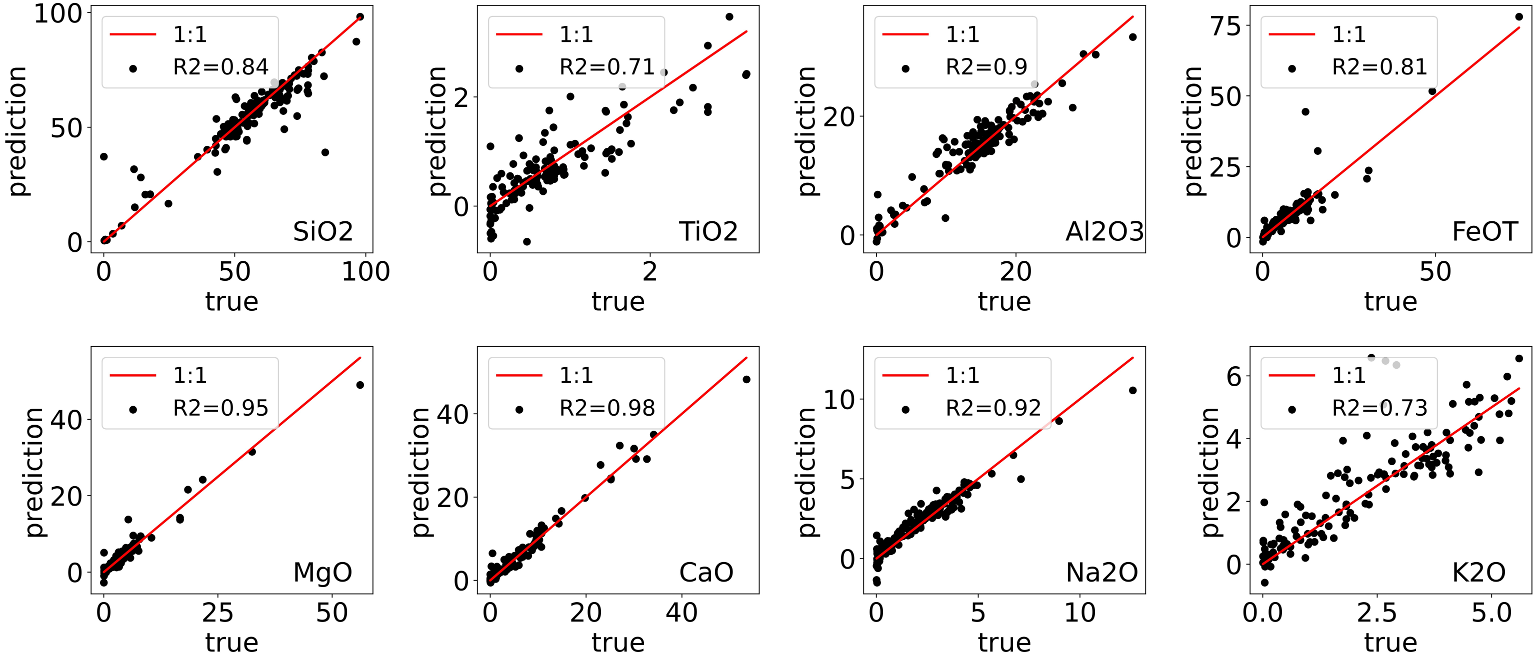
$$v^{(i)} = \sigma(y^T x^{(i)} + b)$$

Accuracy of results

$$R^2 = 1 - \frac{RSS}{TSS}$$

RSS: sum of squares of residuals

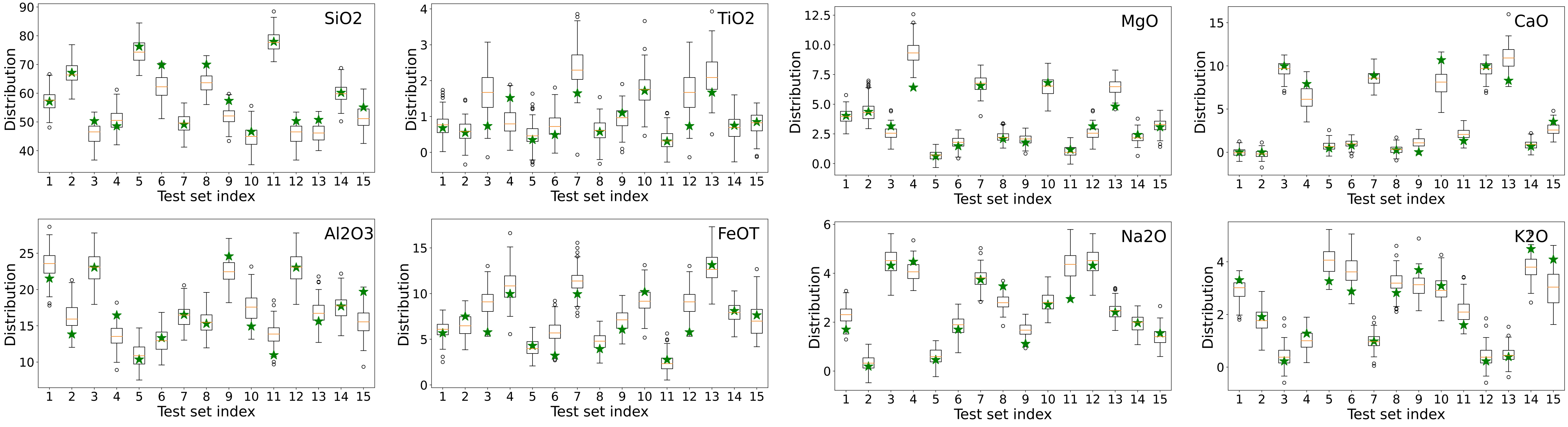
TSS: total sum of squares



Uncertainty quantification via bootstrapping

Bootstrapping: Statistics resampling method that assigns measures of accuracy for any sample estimate

For a new sample $\mathbf{y}_0 \in \mathbb{R}^M$ we get a prediction: $\mathbf{v}_0 = \underbrace{v^{(i)}(\mathbf{y}_0)}_{\text{model uncertainty}} + \underbrace{r(\mathbf{y}_0)}_{\text{data uncertainty}}$



Predictions with uncertainty for 15 random samples

Uncertainty quantification via bootstrapping

Evaluation of prediction intervals

Coverage: The rate at which the *actual* values fall within the range of the prediction interval

Table. Coverage results (95% confidence intervals)

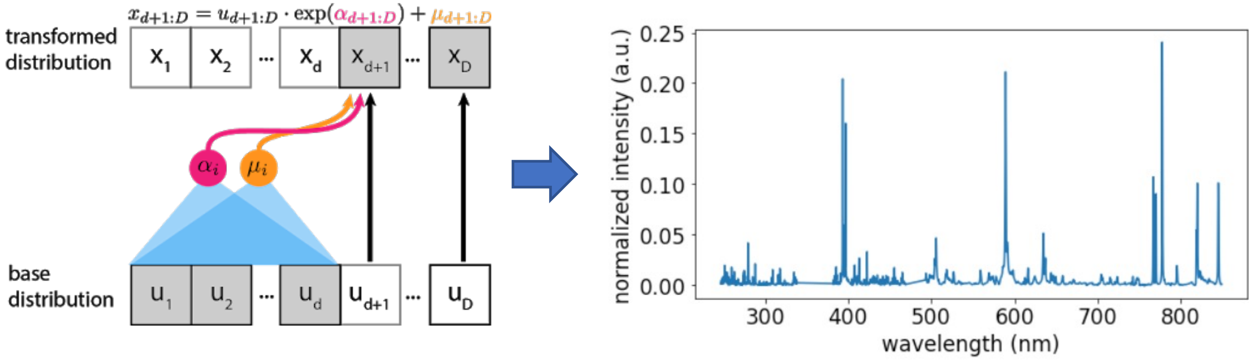
oxide	# of test samples	# of covered samples	Coverage (%)
SiO_2	139	118	86.33
TiO_2	139	137	98.56
Al_2O_3	139	120	86.33
FeO_T	139	120	86.33
MgO	139	120	86.33
CaO	139	134	96.40
Na_2O	139	130	93.53
K_2O	139	125	89.93

Predictions with uncertainty for novel samples

- Samples generated by the normalizing flow model
- Real samples collected on Mars from ChemCam

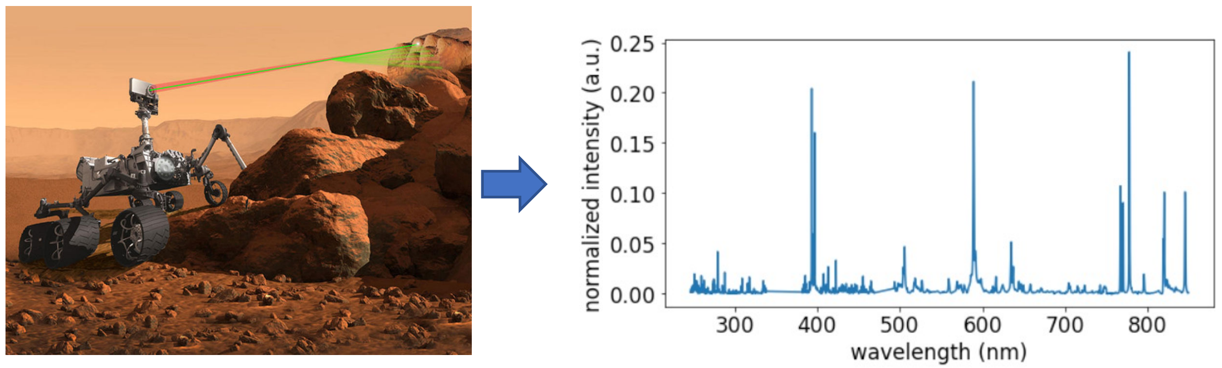
Data generation

Normalizing flow

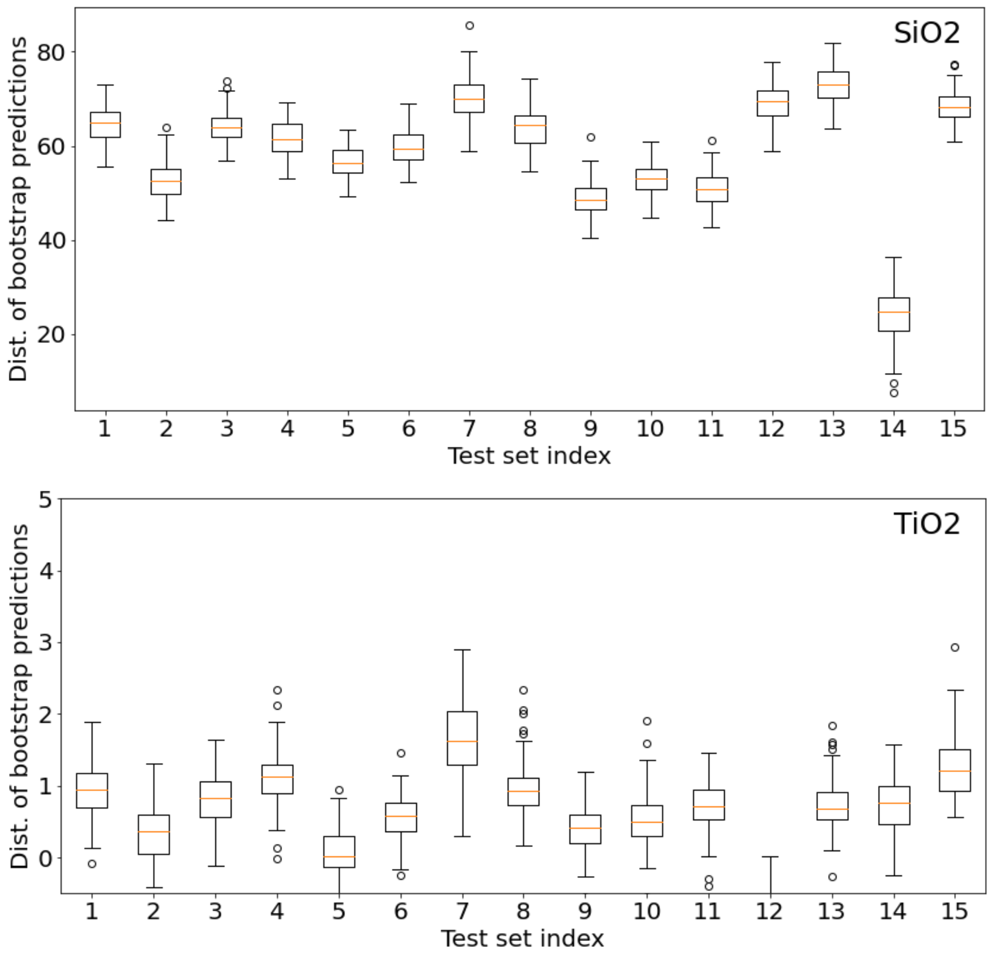


or...

ChemCam



Prediction with uncertainty



Outline

- **Background**: ChemCam data and literature review
- **Motivation**: Uncertainty quantification for ChemCam
- **Methods**: Generative modeling via normalizing flows
- Proposed approach and Results
- Discussion and Conclusions

Conclusions

- Generative modeling can be successfully applied to model real-world data
- Normalizing flow models can be efficiently constructed on latent spaces for fast downstream inference
- Unsupervised and supervised learning can be combined to form an uncertainty quantification framework

Los Alamos NATIONAL LABORATORY

References

1. Wiens et al. "The ChemCam instrument suite on the Mars Science Laboratory (MSL) rover: Body unit and combined system tests." *Space science reviews* 170, no. 1 (2012): 167-227.
2. Forni et al. "Independent component analysis classification of laser induced breakdown spectroscopy spectra." *Spectrochimica Acta Part B: Atomic Spectroscopy* 86 (2013): 31-41.
3. Castorena et al. "Deep spectral CNN for laser induced breakdown spectroscopy." *Spectrochimica Acta Part B: Atomic Spectroscopy* 178 (2021): 106125.
4. Kobyzev et al. "Normalizing flows: An introduction and review of current methods." *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2020).
5. Kumar et al. "Bootstrap prediction intervals in non-parametric regression with applications to anomaly detection." In *Proc. 18th ACM SIGKDD Conf. Knowl. Discovery Data Mining*. 2012.
6. Papamakarios et al. "Normalizing flows for probabilistic modeling and inference." *arXiv preprint arXiv:1912.02762* (2019).
7. Dinh et al. "Density estimation using real nvp." *arXiv preprint arXiv:1605.08803* (2016).

Thank you! ☺